Reference:

[1] Naveen (NNK), "PySpark SQL with Examples," *Spark By {Examples}*, Dec. 19, 2022. https://sparkbyexamples.com/pyspark/pyspark-sql-with-examples/

[2] K. Willems, "PySpark Cheat Sheet: Spark in Python," *Datacamp.com*, Jul. 29, 2021. https://www.datacamp.com/cheat-sheet/pyspark-cheat-sheet-spark-in-python (accessed Mar. 16, 2025).

[3] "pyspark.sql.DataFrameReader.csv — PySpark 3.5.5 documentation," *Apache.org*, 2025. https://spark.apache.org/docs/latest/api/python/reference/pyspark.sql/api/pyspark.sql.DataFrameReader.csv.html?highlight=csv (accessed Mar. 16, 2025).

[4] vergil, "where is the default scheme configuration in hadoop?," *Stack Overflow*, Jun. 10, 2015. https://stackoverflow.com/questions/30749364/where-is-the-default-scheme-configuration-in-hadoop

[5] NNK, "Spark – What is SparkSession Explained - Spark by {Examples}," Mar. 29, 2020. https://sparkbyexamples.com/spark/sparksession-explained-with-examples/

[6] "SQL Cheat Sheet," *GeeksforGeeks*, Mar. 26, 2024. https://www.geeksforgeeks.org/sql-cheat-sheet/

[7] "SQL RANK() Function Explained By Practical Examples," *SQL Tutorial*. https://www.sqltutorial.org/sql-window-functions/sql-rank/

[8] "SQL Cheat Sheet Download PDF it in PDF or PNG Format," *SQL Tutorial*. https://www.sqltutorial.org/sql-cheat-sheet/

[9] "SQL COALESCE Function," *SQL Tutorial*, Feb. 08, 2025. https://www.sqltutorial.org/sql-coalesce/ (accessed Mar. 16, 2025).

[10] "SQL Basics Cheat Sheet," *LearnSQL.com*, Sep. 25, 2020. https://learnsql.com/blog/sql-basics-cheat-sheet/

[11] R. Cotton, "SQL Basics Cheat Sheet," *Datacamp.com*, Mar. 30, 2022. https://www.datacamp.com/cheat-sheet/sql-basics-cheat-sheet

[12] Naveen Nelamali, "PySpark Write to CSV File," *Spark By {Examples}*, Aug. 11, 2022. https://sparkbyexamples.com/pyspark/pyspark-write-dataframe-to-csv-file/ (accessed Mar. 16, 2025).

[13] "pyspark.sql.DataFrameWriter.options — PySpark 3.5.5 documentation," *Apache.org*, 2025. https://spark.apache.org/docs/latest/api/python/reference/pyspark.sql/api/pyspark.sql.DataFrameWriter.options.html (accessed Mar. 16, 2025).

[14] rimmalapudi, "Spark write() Options," *Spark By {Examples}*, Mar. 2023. https://sparkbyexamples.com/spark/spark-write-options/ (accessed Mar. 16, 2025).

[15] "pyspark.sql.SparkSession.stop — PySpark 3.5.5 documentation," *Apache.org*, 2025. https://spark.apache.org/docs/latest/api/python/reference/pyspark.sql/api/pyspark.sql.SparkSession.stop.html?highlight=spark%20stop (accessed Mar. 16, 2025).